

RECEIVED

JUN 07 2002

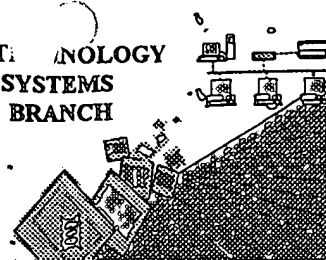
## Raw Sequence Listing Error Summary

TECH CENTER 1600/2900

**ERROR DETECTED      SUGGESTED CORRECTION**SERIAL NUMBER: 09/689,343B

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1      Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file  
    Wrapped Aminos      was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will  
    prevent "wrapping."
- 2      Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3      Misaligned Amino      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers;  
    Numbering      use space characters, instead.
- 4      Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please  
    ensure your subsequent submission is saved in ASCII text.
- 5      Variable Length      Sequence(s)      contain n's or Xaa's representing more than one residue. Per Sequence Rules,  
    each n or Xaa can only represent a single residue. Please present the maximum number of each  
    residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6      PatentIn 2.0      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid  
    "bug"      sequences(s)     . Normally, PatentIn would automatically generate this section from the  
    previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to  
    the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for  
    Artificial or Unknown sequences.
- 7      Skipped Sequences      Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence:  
    (OLD RULES)      (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    (i)      SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
    (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    This sequence is intentionally skipped
- Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8      Skipped Sequences      Sequence(s)      missing. If intentional, please insert the following lines for each skipped sequence.  
    (NEW RULES)      <210> sequence id number  
    <400> sequence id number  
    000
- 9      Use of n's or Xaa's      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
    (NEW RULES)      Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
    In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10    Invalid <213>      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or  
    Response      scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or  
    is Artificial Sequence
- 11      Use of <220>      Sequence(s)      missing the <220> "Feature" and associated numeric identifiers and responses.  
    Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or  
    "Unknown." Please explain source of genetic material in <220> to <223> section.  
    (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12      PatentIn 2.0      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file,  
    "bug"      resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence  
    listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.



RAW SEQUENCE LISTING  
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/689,343B  
Source: 1636  
Date Processed by STIC: 5-30-02

RECEIVED

JUN 07 2002

TECH CENTER 1600/2900

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

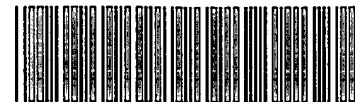
Revised 01/29/2002

Does Not Comply  
Corrected Diskette Needed

RECEIVED

JUN 07 2002

TECH CENTER 1600/2900



1636

## RAW SEQUENCE LISTING

DATE: 05/30/2002

PATENT APPLICATION: US/09/689,343B

TIME: 10:22:31

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\05302002\I689343B.raw

2 <110> APPLICANT: Vaisvila, Romualdus  
 3 Morgan, Richard D.  
 4 Kucera, Rebecca B.  
 5 Claus, Toby B.  
 6 Raleigh, Elisabeth A.  
 8 <120> TITLE OF INVENTION: Method For Cloning And Producing The MseI Restriction  
 Endonuclease  
 10 <130> FILE REFERENCE: NEB-181  
 12 <140> CURRENT APPLICATION NUMBER: US 09/689,343B  
 13 <141> CURRENT FILING DATE: 2000-10-12  
 15 <160> NUMBER OF SEQ ID NOS: 9  
 17 <170> SOFTWARE: PatentIn version 3.1  
 19 <210> SEQ ID NO: 1  
 20 <211> LENGTH: 903  
 21 <212> TYPE: DNA  
 22 <213> ORGANISM: Micrococcus sp.  
 24 <220> FEATURE:  
 25 <221> NAME/KEY: CDS  
 26 <222> LOCATION: (1)..(900)  
 27 <223> OTHER INFORMATION:  
 29 <400> SEQUENCE: 1  
 30 atg cct atc tcg acc gtc tgg acg ccg gac gga gac gac ctc atc gtg 48  
 31 Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val  
 32 1 5 10 15  
 34 gag gcg gac aac ctc gat ttc att caa acg ctc ccc gac gcg agc ttc 96  
 35 Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe  
 36 20 25 30  
 38 cga atg atc tac atc gat ccg ccg ttc aac aca ggg cga acg cag ccg 144  
 39 Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg  
 40 35 40 45  
 42 ctt cag tcg ctc aag acg acc cgc tcg gtc aca ggg tcg cga gtc ggc 192  
 43 Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly  
 44 50 55 60  
 46 ttc aaa ggc cag acg tac gac acg gtc aag agc act ctg cac tcg tat 240  
 47 Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr  
 48 65 70 75 80  
 50 gac gac gct ttc acc gac tat tgg tcg ttc ctc gaa ccg cgt ctc ctg 288  
 51 Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu  
 52 85 90 95  
 54 gag gct tgg cgg ttg ctc acc cct gac ggc gcg ctc tat ctt cat ctg 336  
 55 Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu  
 56 100 105 110  
 58 gat tac cgc gag gtt cac tac gcc aag gtc gtc ctc gac gcg atg ttc 384  
 59 Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe

## RAW SEQUENCE LISTING

DATE: 05/30/2002

PATENT APPLICATION: US/09/689,343B

TIME: 10:22:31

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\05302002\I689343B.raw

```

60          115          120          125
62 gga cgc gaa agc ttc ctg aac gag ctg atc tgg gcg tac gac tac ggc      432
63 Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly
64          130          135          140
66 gcg cgc tcg aag agc aag tgg ccc acc aag cac gac aac atc ctc gtg      480
67 Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val
68 145          150          155          160
70 tat gtg aag gac ccg aac aac tac gtc tgg aac ggt cag gat gta gat      528
71 Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp
72          165          170          175
74 cgc gag ccc tac atg gcg ccc ggg ctc gtt aca ccc gag aag gta gcg      576
75 Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala
76          180          185          190
78 ctt ggc aag ctg ccc acc gac gtc tgg tgg cac aca atc gtt ccg cct      624
79 Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro
80          195          200          205
82 gcg agc aaa gag cgc acc ggg tac gcg aca cag aag ccg gtc ggc atc      672
83 Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile
84          210          215          220
86 atc cgt cgc atg att cag gcg agc agc aat gaa ggc gac tgg gtt ctg      720
87 Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu
88 225          230          235          240
90 gat ttc ttc gct ggt agt ggg acg acc ggc gcc gcg gcc cgc cag ctc      768
91 Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu
92          245          250          255
94 gga cgc cgt ttt gtg ctc gta gac gtc aac cca gaa gca atc gcg gta      816
95 Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val
96          260          265          270
98 atg gca aaa cgg ttg gat gac ggg gca ttg gac acc agc gtg acg atc      864
99 Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile
100          275          280          285
102 gtg cag act ccc cag agt gac cca cga acc gac gga tga      903
103 Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly
104          290          295          300
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 300
109 <212> TYPE: PRT
110 <213> ORGANISM: Micrococcus sp.
112 <400> SEQUENCE: 2
114 Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val
115 1          5          10          15
118 Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe
119          20          25          30
122 Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg
123          35          40          45
126 Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly
127          50          55          60
130 Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr
131 65          70          75          80

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/689,343B

DATE: 05/30/2002

TIME: 10:22:32

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\05302002\I689343B.raw

```

134 Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu
135           85           90           95
138 Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu
139           100           105           110
142 Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe
143           115           120           125
146 Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly
147           130           135           140
150 Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val
151 145           150           155           160
154 Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp
155           165           170           175
158 Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala
159           180           185           190
162 Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro
163           195           200           205
166 Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile
167           210           215           220
170 Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu
171 225           230           235           240
174 Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu
175           245           250           255
178 Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val
179           260           265           270
182 Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile
183           275           280           285
186 Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly
187           290           295           300
190 <210> SEQ ID NO: 3
191 <211> LENGTH: 1236
192 <212> TYPE: DNA
193 <213> ORGANISM: Environmental DNA
195 <220> FEATURE:
196 <221> NAME/KEY: CDS
197 <222> LOCATION: (1)..(1233)
198 <223> OTHER INFORMATION:
200 <220> FEATURE:
201 <221> NAME/KEY: misc_feature
202 <222> LOCATION: (198)..(198)
203 <223> OTHER INFORMATION: Xaa = any amino acid
205 <220> FEATURE:
206 <221> NAME/KEY: misc_feature
207 <222> LOCATION: (594)..(594)
208 <223> OTHER INFORMATION: N= G, A, C or T
210 <400> SEQUENCE: 3
211 atg cct aca ctg gat tgg ccc ggt aaa cag tta agc ttc cca cca gct      48
212 Met Pro Thr Leu Asp Trp Pro Gly Lys Gln Leu Ser Phe Pro Pro Ala
213 1           5           10           15
215 acc tcc ttg cat ctg gag agt gtg gtc act gag gga gcg gag tca ccg      96

```

See item # 10 on  
Error Summary sheet  
↑↑

The type of errors shown exist throughout  
the Sequence Listing. Please check subsequent  
sequences for similar errors.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/689,343B

DATE: 05/30/2002

TIME: 10:22:32

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\05302002\I689343B.raw

216	Thr	Ser	Leu	His	Leu	Glu	Ser	Val	Val	Thr	Glu	Gly	Ala	Glu	Ser	Pro	
217				20					25					30			
219	cct	aat	cgt	ctg	att	tgg	gcg	gac	aac	ctg	ccg	cta	atg	gta	gat	ttg	144
220	Pro	Asn	Arg	Leu	Ile	Trp	Ala	Asp	Asn	Leu	Pro	Leu	Met	Val	Asp	Leu	
221			35					40					45				
223	ttg	gcc	gaa	tat	gaa	ggg	aaa	atc	gat	ctg	atc	tac	gcc	gat	ccc	cct	192
224	Leu	Ala	Glu	Tyr	Glu	Gly	Lys	Ile	Asp	Leu	Ile	Tyr	Ala	Asp	Pro	Pro	
225		50				55					60						
227	ttt	ttt	acg	gat	cgt	act	tat	gcg	gcg	cga	att	ggt	cat	ggg	gag	gat	240
228	Phe	Phe	Thr	Asp	Arg	Thr	Tyr	Ala	Ala	Arg	Ile	Gly	His	Gly	Glu	Asp	
229	65					70				75					80		
231	tcg	cgt	cgt	cca	caa	acc	tgg	cag	ctt	gca	gaa	gga	tat	acg	gac	gag	288
232	Ser	Arg	Arg	Pro	Gln	Thr	Trp	Gln	Leu	Ala	Glu	Gly	Tyr	Thr	Asp	Glu	
233				85					90					95			
235	tgg	aag	gat	tta	gat	gaa	tac	ctg	gac	ttc	ctt	tat	cca	cgc	ctg	gta	336
236	Trp	Lys	Asp	Leu	Asp	Glu	Tyr	Leu	Asp	Phe	Leu	Tyr	Pro	Arg	Leu	Val	
237			100					105					110				
239	ctg	atg	tat	cga	ctg	ctg	gca	cca	cac	gga	acg	ctc	tac	ttg	cac	ctg	384
240	Leu	Met	Tyr	Arg	Leu	Leu	Ala	Pro	His	Gly	Thr	Leu	Tyr	Leu	His	Leu	
241			115				120						125				
243	gac	tgg	cac	gcc	aat	gcc	tac	gta	cgt	gta	ctg	ctt	gat	gag	atc	ttc	432
244	Asp	Trp	His	Ala	Asn	Ala	Tyr	Val	Arg	Val	Leu	Leu	Asp	Glu	Ile	Phe	
245		130				135					140						
247	ggg	cga	cag	cgg	ttt	ctc	aac	gag	atc	gtc	tgg	atc	tat	cac	ggc	ccc	480
248	Gly	Arg	Gln	Arg	Phe	Leu	Asn	Glu	Ile	Val	Trp	Ile	Tyr	His	Gly	Pro	
249	145				150					155					160		
251	tca	gcc	atc	cga	cgc	gcc	ttc	aag	cgc	aaa	cat	gat	acc	atc	ttg	gtt	528
252	Ser	Ala	Ile	Arg	Arg	Ala	Phe	Lys	Arg	Lys	His	Asp	Thr	Ile	Leu	Val	
253				165					170					175			
255	tat	gtg	aaa	ggt	gaa	aac	tat	aca	ttc	aat	gcg	gat	gcg	gtt	cgt	caa	576
256	Tyr	Val	Lys	Gly	Glu	Asn	Tyr	Thr	Phe	Asn	Ala	Asp	Ala	Val	Arg	Gln	
257			180					185					190				
W--> 259	cct	tac	cat	ccg	agc	acn	cat	aag	acc	ttc	gct	tcc	tcc	ccg	aag	gcc	624
W--> 260	Pro	Tyr	His	Pro	Ser	Xaa	His	Lys	Thr	Phe	Ala	Ser	Ser	Pro	Lys	Ala	
261			195				200					205					
263	ggc	ttt	ggt	aag	gtg	ccg	gat	ctg	cag	cgc	ggc	aaa	gtg	ccc	gaa	gac	672
264	Gly	Phe	Gly	Lys	Val	Pro	Asp	Leu	Gln	Arg	Gly	Lys	Val	Pro	Glu	Asp	
265		210				215					220						
267	tgg	tgg	tat	ttt	ccg	gtc	gtg	gcc	cgt	cta	cac	cga	gaa	cgg	agc	ggc	720
268	Trp	Trp	Tyr	Phe	Pro	Val	Val	Ala	Arg	Leu	His	Arg	Glu	Arg	Ser	Gly	
269	225				230				235					240			
271	tat	ccg	act	caa	aag	cct	caa	gcc	ttg	ctg	gag	cgg	atc	ctg	ctg	gcc	768
272	Tyr	Pro	Thr	Gln	Lys	Pro	Gln	Ala	Leu	Leu	Glu	Arg	Ile	Leu	Leu	Ala	
273				245				250					255				
275	tcc	tcg	aac	gca	ggc	gat	ctg	gtg	gca	gac	ttc	ttc	tgc	ggc	tca	ggg	816
276	Ser	Ser	Asn	Ala	Gly	Asp	Leu	Val	Ala	Asp	Phe	Phe	Cys	Gly	Ser	Gly	
277			260				265					270					
279	aca	acc	gct	gtg	gtg	gca	gcc	cgt	ctg	gga	cgg	cgc	ttc	ctg	gtc	aac	864
280	Thr	Thr	Ala	Val	Val	Ala	Ala	Arg	Leu	Gly	Arg	Arg	Phe	Leu	Val	Asn	

## RAW SEQUENCE LISTING

DATE: 05/30/2002

PATENT APPLICATION: US/09/689,343B

TIME: 10:22:32

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\05302002\I689343B.raw

```

281          275          280          285
283 gat gca agc tgg cgc gcc gtt cat gtg aca cgc aca cgc ttg cta cgc      912
284 Asp Ala Ser Trp Arg Ala Val His Val Thr Arg Thr Arg Leu Leu Arg
285          290          295          300
287 gag gga gta agt ttc act ttt gaa cgc cag gaa act ttt act cta cct      960
288 Glu Gly Val Ser Phe Thr Phe Glu Arg Gln Glu Thr Phe Thr Leu Pro
289 305          310          315          320
291 atc cag cca ctt cca cca gat tgg ttg atc atc gcc gag gag cag att      1008
292 Ile Gln Pro Leu Pro Pro Asp Trp Leu Ile Ile Ala Glu Glu Gln Ile
293          325          330          335
295 cgc ctc caa gca ccc ttt ctc gta gat ttt tgg gaa gtg gac gat caa      1056
296 Arg Leu Gln Ala Pro Phe Leu Val Asp Phe Trp Glu Val Asp Asp Gln
297          340          345          350
299 tgg gat ggc aaa atc ttc cgc agc cgt cat caa ggc tta cgc tcc cgc      1104
300 Trp Asp Gly Lys Ile Phe Arg Ser Arg His Gln Gly Leu Arg Ser Arg
301          355          360          365
303 ctt cag gag cag gcg ccg ctc tct cta cca ttg acc ggg aat gga ctg      1152
304 Leu Gln Glu Gln Ala Pro Leu Ser Leu Pro Leu Thr Gly Asn Gly Leu
305          370          375          380
307 ttg tgt gta cgg gta gtg agc cgt gaa ggg gaa tac tat gag ttc aca      1200
308 Leu Cys Val Arg Val Val Ser Arg Glu Gly Glu Tyr Tyr Glu Phe Thr
309 385          390          395          400
311 ggt cga gcc gat agc cct cac ccc gta tcg ttt tga      1236
312 Gly Arg Ala Asp Ser Pro His Pro Val Ser Phe
313          405          410
316 <210> SEQ ID NO: 4
317 <211> LENGTH: 411
318 <212> TYPE: PRT
319 <213> ORGANISM: Environmental DNA
321 <220> FEATURE:
322 <221> NAME/KEY: misc_feature
323 <222> LOCATION: (198)..(198)
324 <223> OTHER INFORMATION: Xaa = any amino acid
328 <400> SEQUENCE: 4
330 Met Pro Thr Leu Asp Trp Pro Gly Lys Gln Leu Ser Phe Pro Pro Ala
331 1          5          10          15
334 Thr Ser Leu His Leu Glu Ser Val Val Thr Glu Gly Ala Glu Ser Pro
335          20          25          30
338 Pro Asn Arg Leu Ile Trp Ala Asp Asn Leu Pro Leu Met Val Asp Leu
339          35          40          45
342 Leu Ala Glu Tyr Glu Gly Lys Ile Asp Leu Ile Tyr Ala Asp Pro Pro
343          50          55          60
346 Phe Phe Thr Asp Arg Thr Tyr Ala Ala Arg Ile Gly His Gly Glu Asp
347 65          70          75          80
350 Ser Arg Arg Pro Gln Thr Trp Gln Leu Ala Glu Gly Tyr Thr Asp Glu
351          85          90          95
354 Trp Lys Asp Leu Asp Glu Tyr Leu Asp Phe Leu Tyr Pro Arg Leu Val
355          100          105          110
358 Leu Met Tyr Arg Leu Leu Ala Pro His Gly Thr Leu Tyr Leu His Leu

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/689,343B

DATE: 05/30/2002  
TIME: 10:22:33

Input Set : A:\NEB-181.ST25.txt  
Output Set: N:\CRF3\05302002\I689343B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 594  
Seq#:3; Xaa Pos. 198  
Seq#:4; Xaa Pos. 198



VERIFICATION SUMMARY

PATENT APPLICATION: US/09/689,343B

DATE: 05/30/2002

TIME: 10:22:33

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\05302002\I689343B.raw

L:259 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:576

L:260 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:624

L:378 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:192